THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 15, Issue of April 12, pp. 13138–13147, 2002

Printed in U.S.A.

Characterization and Expression Analysis of *Staphylococcus aureus* Pathogenicity Island 3

IMPLICATIONS FOR THE EVOLUTION OF STAPHYLOCOCCAL PATHOGENICITY ISLANDS*

Received for publication, December 6, 2001, and in revised form, January 3, 2002
Published, JBC Papers in Press, January 30, 2002, DOI 10.1074/jbc.M111661200

Jeremy M. Yarwood‡§, John K. McCormick‡ ‡‡, Michael L. Paustian¶, Paul M. Orwin‡‖, Vivek Kapur¶, and Patrick M. Schlievert‡**

From the ‡Department of Microbiology, University of Minnesota Medical School, Minneapolis, Minnesota 55455, and ¶Department of Veterinary Pathobiology and Biomedical Genomics Center, University of Minnesota, St. Paul, Minnesota 55108

We describe the complete sequence of the 15.9-kb staphylococcal pathogenicity island 3 encoding staphylococcal enterotoxin serotypes B, K, and Q. The island, which meets the generally accepted definition of pathogenicity islands, contains 24 open reading frames potentially encoding proteins of more than 50 amino acids, including an apparently functional integrase. The element is bordered by two 17-bp direct repeats identical to those found flanking staphylococcal pathogenicity island 1. The island has extensive regions of homology to previously described pathogenicity islands, particularly staphylococcal pathogenicity islands 1 and bov. The expression of 22 of the 24 open reading frames contained on staphylococcal pathogenicity island 3 was detected either *in vitro* during growth in a laboratory medium or serum or *in vivo* in a rabbit model of toxic shock syndrome using DNA microarrays. The effect of oxygen tension on staphylococcal pathogenicity island 3 gene expression was also examined. By comparison with the known staphylococcal pathogenicity islands in the context of gene expression described here, we propose a model of pathogenicity island origin and evolution involving specialized transduction events and addition, deletion, or recombination of pathogenicity island "modules."

Staphylococcus aureus is a leading etiologic agent of both nosocomial and community-acquired infections worldwide. These infections range from fairly benign cutaneous infections, such as furuncles, to potentially fatal diseases, including endocarditis and toxic shock syndrome (TSS)[1] (reviewed in Ref. 1). Its ability to cause this range of disease is due in part to its elaboration of a vast array of both cell surface-associated and secreted virulence factors. Among the secreted factors are the pyrogenic toxin superantigens that have the ability to activate large populations (10–50%) of T lymphocytes in a manner specific to the variable region of the β-chain of the T-cell receptor (2). The ensuing massive cytokine release results in the symptoms of TSS, including fever, hypotension, rash, vomiting, diarrhea, multiple organ failure, disseminated intravascular coagulation, and desquamation. The staphylococcal enterotoxins (SEs), members of the superantigen family, are associated with both TSS and food poisoning and have proven emetic activities that appear to be separable from their superantigenic activity (3).

Most, if not all, staphylococcal superantigens are encoded by accessory genetic elements that are either mobile or appear to have been mobile at one time (reviewed in Ref. 4). These identified elements include plasmids, transposons, prophages, and the pathogenicity islands. The staphylococcal pathogenicity islands (SaPIs), of which five have, until recently, been described (SaPI1–4 and SaPIbov), are the first clearly defined pathogenicity islands in Gram-positive bacteria, and each encodes one or more of the staphylococcal superantigens (reviewed in Ref. 5). SEB, SEC, SEK, SEL, SEQ, and toxic shock syndrome toxin 1 (TSST-1) are known to be encoded by one or more of these phage-related elements (reviewed in Ref. 4). More recently, Kuroda *et al.* (6) identified six novel pathogenicity islands in the complete genomes of two *S. aureus* strains, N315 and Mu50, including three that carried *tstH* (encoding TSST-1), two that carried clusters of staphylococcal exotoxin-like proteins, and one, present in both strains, that carried clusters of serine proteases and enterotoxins.

These genomic loci meet the generally accepted requirements of the pathogenicity island subgroup of "genomic islands" as defined previously (7, 8). They are present in the genomes of many staphylococci but absent from closely related strains, they are relatively large genomic fragments (>15 kb), they differ in GC content from the rest of the chromosome, they are flanked by direct repeats likely generated upon insertion of the elements into the genome, some are associated with tRNA loci, and they possess genes coding for genetic mobility, including conserved integrases.

The prototypical staphylococcal pathogenicity island, SaPI1, was identified and characterized by Lindsay *et al.* (9) as the

* This work was supported in part by National Institutes of Health Grant AI22159 (to P. M. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF410775.

§ Supported by a Howard Hughes Predoctoral Fellowship in the Biological Sciences.

‖ Present address: Dept. of Environmental Science and Engineering, California Institute of Technology, 209 Keck Laboratories, M/C 138-78, Pasadena, CA 91125-7800.

** To whom correspondence should be addressed: Dept. of Microbiology, University of Minnesota Medical School, MMC 196, 420 Delaware St. SE, Minneapolis, MN 55455. Tel.: 612-624-9471; Fax: 612-626-0623; E-mail: pats@lenti.med.umn.edu.

‡‡ Present address: The Lawson Health Research Inst., The University of Western Ontario, Grosvenor Campus, 268 Grovenor Street, Rm. H-323, London, Ontario, Canada N6A4V2.

[1] The abbreviations used are: TSS, toxic shock syndrome; SaPI, staphylococcal pathogenicity island; TSST-1, toxic shock syndrome toxin 1; ORF, open reading frame; TH, Todd-Hewitt; SE, staphylococcal enterotoxin.

genetic element encoding TSST-1, the only superantigen to be associated with nearly all cases of menstrual TSS. SaPI1 is 15.2 kb in length, flanked by a 17-nucleotide direct repeat, contains a functional integrase (*int*) gene, and is located near the *tyrB* locus in strain RN4282. It also appears to encode a second superantigen, SEK, and part of a third superantigen, SEQ. Mobility of SaPI1 has been demonstrated only in the presence of a helper phage, such as 80α. Ruzin *et al.* (10) have demonstrated that SaPI1 appears to parasitize excision, replication, and encapsidation functions of phage 80α in a relationship that is similar to that between coliphages P4 and P2. During the growth of phage 80α, SaPI1 excises from its unique chromosomal insertion site, att_c, replicates in the linear form, interferes with phage growth, and is encapsidated into specialized phage heads. Upon transduction to a recipient organism, SaPI1 integrates by the classical Campbell mechanism into the att_c site for which the SaPI1-coded integrase is necessary. Because islands with different *att* sites appear to have dissimilar integrases (6), it may well be that the integrase carried by the island determines the integration site in the genome.

Existence of these toxin genes on mobile genetic elements implies their transfer between staphylococcal strains as well as other bacterial species by horizontal transfer. Furthermore, these elements are not uniformly distributed among clinical isolates. Thus, these mobile elements likely have played and continue to play an integral role in the evolution of *S. aureus* as a species and as a pathogen. Indeed, recent evidence supports the hypothesis that virulence traits are spread by horizontal transfer, particularly in nosocomial infections, and that the presence of accessory genetic elements within a strain may affect the acquisition and loss of other mobile genetic elements (11). These islands may also form the basis for toxin gene exclusion. For instance, in testing thousands of strains, our laboratory has never identified a clinical isolate that produced both TSST-1 and SEB. It has been determined that these toxins are encoded by different pathogenicity islands that appear to exclude each other from their respective integration sites.

However, despite the identification of numerous pathogenicity islands and their likely importance in the evolution of *Staphylococcus* as a pathogen, the origin and functions of pathogenicity islands remain areas with little investigation. All of the SaPIs have multiple open reading frames (ORFs), many of which have no identifiable homologs. To this point, the issues of whether or not these ORFs are expressed and whether their function might be in the regulation of island-associated superantigens or only in the maintenance and transfer of the islands have not been addressed. In this study, we report the complete sequence and map of SaPI3 encoding SEB and the more recently identified enterotoxins, SEK (12) and SEQ.[2] Furthermore, we describe for the first time the expression of the numerous genes contained on a staphylococcal pathogenicity island using DNA microarray technology. We then discuss the implications for the evolution of the staphylococcal pathogenicity islands.

MATERIALS AND METHODS

Strains—COL is a prototypical methicillin-resistant isolate of *S. aureus* that is currently being sequenced by The Institute for Genomic Research. MN NJ is a methicillin-sensitive isolate of *S. aureus* from a case of nonmenstrual TSS in which our laboratory has identified two novel superantigens, SEK (12) and SEQ.[2] *S. aureus* MN8, also a methicillin-sensitive isolate of *S. aureus*, was used as a source of genomic template for amplification of virulence gene probes for the DNA microarrays and was isolated before 1980 from a case of menstrual TSS (13).

Sequencing—Preliminary sequence data for the *S. aureus* strain COL was obtained from The Institute for Genomic Research website (www.tigr.org). A primer walking-based approach was used to amplify by PCR and sequence ~3.4 kb that spanned the unsequenced loci of SaPI3 in the COL genomic sequence data base. Automated sequencing using an ABI model 377 was performed with the assistance of the Advanced Genetic Analysis Center (University of Minnesota, St. Paul, MN).

Growth of S. aureus in vitro—50-ml cultures of *S. aureus* MN NJ were grown aerobically with shaking at 37 °C in either Todd-Hewitt (TH) broth (BC PharMingen) or rabbit serum (Invitrogen). Two independent cultures in each medium were grown in parallel, and samples were removed at the exponential, postexponential, and stationary phases of growth (2, 3, and 8 h after inoculation with an initial cell density of $A_{600\ nm}$ = 0.1). Expression of SaPI3 ORFs and other virulence-associated genes was quantified using DNA microarrays.

For cultures exposed to altered oxygen levels, 1 ml of TH broth was inoculated with *S. aureus* MN NJ from an overnight culture to an initial cell density of $A_{600\ nm}$ = 0.1 in a 35 × 12-mm-diameter polystyrene Petri dish (Nunc, Roskilde, Denmark). All cultures were placed into sealed, humidified Plexiglass cell culture chambers (20 × 26 × 7.5 cm, internal dimensions; Mishell-Dutton) (14), flushed with gas mixtures containing either 1% oxygen (v/v) or 21% oxygen (v/v) balanced with nitrogen and 7% carbon dioxide (Praxair, St. Louis, MO), and sealed. Chambers were then incubated at 37 °C with orbital shaking (~125 rpm). Samples for the exponential, postexponential, and stationary phases of growth were removed at ~2, 3, and 8 h after inoculation, respectively, and expression of SaPI3 genes was quantified using DNA microarrays. We have previously determined that there is no significant difference in growth parameters (cell densities or timing of entry into the growth phases of interest) between cultures grown in 1% oxygen and 21% oxygen (v/v) balanced with nitrogen and 7% carbon dioxide (15, 16).

Immunization of Dutch-belted Rabbits—Two Dutch-belted rabbits were immunized with SEB, which is made in high concentrations (>10 μg/ml) by MN NJ grown *in vitro*. Rabbits were immunized by three subcutaneous injections at 2-week intervals, with each injection containing 25 μg of purified SEB resuspended in 0.5 ml of phosphate-buffered saline and emulsified in 0.5 ml of incomplete Freund's adjuvant. Development of antibody to SEB was determined by enzyme-linked immunosorbent assay of serum samples taken 1 week after the final immunization. The two rabbits developed anti-SEB (IgG) titers of 1:5,120 and 1:10,240, respectively, as compared with preimmune titers of <1:20.

Subcutaneous Infection Model—Sterilized perforated hollow polyethylene golf balls were implanted subcutaneously in four Dutch-belted rabbits (17). Implantation of the polyethylene balls and subsequent healing created transudate-filled cavities in the rabbits with volumes of ~15 ml that contained few host cells, thus enabling the preparation of staphylococcal RNA relatively free of contaminating host RNA. 6 weeks after implantation of the polyethylene balls, ~10^{10} colony-forming units of *S. aureus* MN NJ grown in TH medium were collected by centrifugation from the late exponential phase of growth (cell density of cultures was 6.7 × 10^8 colony-forming units/ml), resuspended in 2 ml of phosphate-buffered saline, and injected into the implanted polyethylene balls. Samples were removed from the inoculum culture before centrifugation for use in expression analysis by DNA microarrays. 2 ml of transudate containing *S. aureus* were then removed from the infection chambers at the indicated times after inoculation using a sterile syringe, *S. aureus* was enumerated by plating, and expression of SaPI3 genes was quantified using DNA microarrays.

RNA Preparation and DNA Microarrays—Analysis of staphylococcal gene expression *in vitro* and *in vivo* using DNA microarrays was performed as described elsewhere (www.agac.umn.edu/microarray/protocols/protocols.htm). In brief, a library of targets representing 68 genes from *S. aureus* MN NJ and MN8 was constructed with primers designed to amplify fragments of ~300 bp of each gene from genomic DNA. Two successive rounds of PCR were performed to minimize genomic DNA contamination in the amplification products, and the final 100-μl reactions were checked for quality on agarose gels and purified with the QIAquick PCR Purification Kit (Qiagen, Valencia, CA). The purified products were printed in triplicate using a Total Array System robot (BioRobotics, Boston, MA). Cell pellets from centrifuged samples of *S. aureus* cultures were flash-frozen in liquid nitrogen. Total RNA was prepared using the RNeasy Mini Kit (Qiagen) according to the manufacturer's directions. DNA was removed from the RNA preparations using the RNase-free DNase Set (Qiagen) according to the manufacturer's directions. cDNA prepared from RNA from *S. aureus* cultures to be compared was labeled with either Cy3 or Cy5 fluorescent dye (Amer-

[2] P. M. Orwin, D. Y. M. Leung, H. L. Donahue, G. A. Bohach, and P. M. Schlievert, submitted for publication.

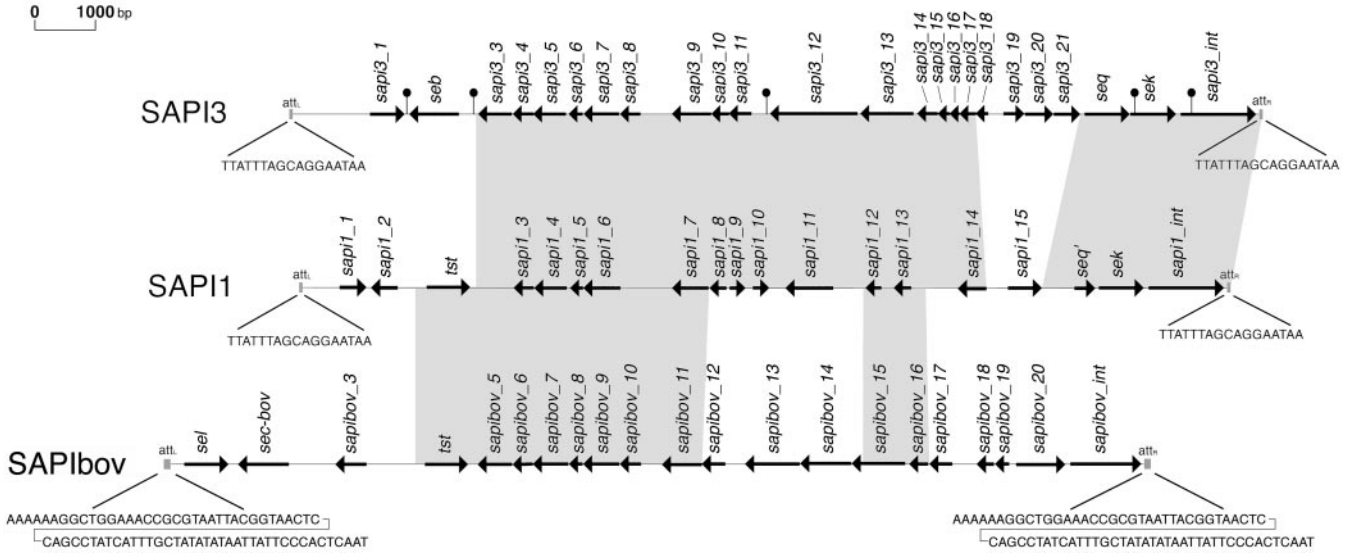

FIG. 1. **Comparative map of pathogenicity islands 3 (*SaPI3*), 1 (*SaPI1*), and bov (*SaPIbov*).** *Arrows* represent the location and orientation of open reading frames or previously described genes greater than 50 amino acids in length. Regions of sequence homology greater than 92% between the islands are indicated by *shading*. Genes encoding enterotoxins B (*seb*), K (*sek*), Q (*seq*), and the bovine variant of C (*sec-bov*) are indicated. *Seq′* potentially encodes a truncated enterotoxin L. *sapi3_1* (*ear*) encodes a putative β-lactamase-like protein. Also shown are the likely integrase genes (*int*) as well as the attachment (*att*) sequences. Stem-loop structures that may serve as rho-independent transcription terminators are indicated for SaPI3 (predicted using DNA Strider version 1.2).

sham Biosciences) and competitively hybridized with the printed microarrays. Images of the hybridized arrays were obtained with a Scanarray 5000 microarray scanner (GSI Lumonics, Watertown, MA). One independent hybridization (on triplicate arrays) was conducted for each of two independent experiments. Fluorescence intensities for individual spots were normalized based on the total intensity of fluorescence in the Cy3 and Cy5 channels. Fluorescence intensity was determined as the average intensity of the triplicate spots for each gene. Total fluorescence for each gene was normalized between arrays for independent experiments, the data were combined from both experiments, and statistical significance was determined using Student's *t* test to compare expression data from the two growth conditions of interest. SaPI3 ORFs were determined as being expressed if the fluorescence intensity was at least twice that of background levels established using negative controls (probes for genes not expressed by strains MN NJ and COL) and if fluorescence was detected in each of the triplicate arrays for each independent experiment. To account for possible bias in labeling of cDNA by either Cy3 or Cy5, dye labeling was reversed in the second independent experiment for each of several experimental conditions. No dye bias was detected. Clustering based on similarity of expression profiles and visualization were performed using the software program Spotfire DecisionSite 6.1 (www.spotfire.com). Similarities between expression profiles of individual genes in all 11 experimental conditions were calculated using the "Euclidean distance" method.

RESULTS

Identification and Characterization of SaPI3—Data base searches of the unfinished *S. aureus* COL genome (www.tigr.org) revealed the presence of large segments homologous to SaPI1. Because COL does not produce TSST-1, we hypothesized that this sequence comprised a novel pathogenicity island, which we termed SaPI3. Using the SaPI1 sequence and SaPI3 partial sequence as guides, primers were designed to complete the sequence of SaPI3 in the COL strain. In all, a PCR and primer walking-based approach was used to sequence ~3.4 kb of the 15,936-bp SaPI3. SaPI3 was determined to have 24 ORFs potentially encoding proteins over 50 amino acids in length, 3 of which encoded staphylococcal enterotoxin serotypes B, K, and Q, and many of which have homologs in SaPI1 and SaPIbov (Fig. 1; Table I). We were also able to identify the presence of SaPI3 in the clinical isolate *S. aureus* MN NJ, a known SEB producer, by PCR analysis and sequencing of the same three regions as in strain COL (data not shown). MN NJ is an isolate from a case of nonmenstrual TSS in which our laboratory has described the presence of two novel enterotoxins, SEK (12) and SEQ.[2] The repeated 17-nucleotide sequences flanking SaPI3 were identical to the *att* sites of SaPI1 (5′-TTATTTAGCAGGATAA-3′) and thus might form the basis of pathogenicity island exclusion (*i.e.* the lack of TSST-1 and SEB in the same clinical isolates). In general, the identified SaPI3 genes form two apparent transcriptional blocks with ORFs 2–18 (including SEB) oriented toward the left of the island, whereas ORFs 19–24 (including SEK and SEQ) are oriented to the right (Fig. 1). The overall GC content of SaPI3 was 31.4%, somewhat lower than the 32.8–32.9% found for the whole genome of *S. aureus* (6).

We have identified genes in SaPI3 according to the following nomenclature: *pathogenicity island_ORF number (variant)*. Thus, the ninth ORF in SaPI3 is identified by the gene name *sapi3_9*. A mutant of this gene might be designated as *sapi3_9(1)*. Sequential genes are identified using a hyphen (*e.g.* sapi3_10-15 is used to identify all SaPI3 ORFs from 10 through 15). Upon determination of the gene's function, the name will be altered to reflect that function, such as *sapi3_int*. (Exotoxins on the islands are identified according to their own nomenclature system.) If additional genes are identified on the island subsequent to its initial sequencing, those genes will be numbered sequential to those already identified, rather than renumbering all of the genes on the island. We have implemented this nomenclature system in our laboratory to promote systematic identification of SaPI genes, preclude confusion regarding identically numbered ORFs on different islands, and allow unambiguous assignation of expression data to SaPI genes.

A comparison of SaPI3, SaPI1, and SaPIbov is shown in Fig. 1, and the corresponding ORFs are identified in Table I. The overall length of the three SaPIs is similar, although SaPIbov is larger on the 5′ end by 1915 bp. SaPIbov also has a different *att* site than SaPI1 and SaPI3. Two core regions of high (>92% identity) homology between all three islands were identified (Fig. 1). The first core region includes nucleotides 2974–6709 of SaPI1, 3113–6929 of SaPI3, and 5100–8745 of SaPIbov. Within this core region, SaPI1 has an additional 100 bp not present in SaPIbov (nucleotides 5909–6009 in SaPI1), whereas

TABLE I
Description of SaPI3 ORFs

Name	Length	Predicted molecular mass[a]	Corresponding ORF in SaPI1[b]	Corresponding ORF in SaPIbov	Description (homolog)	Expression detected *in vitro*	Expression detected *in vivo*
	(amino acids)	*(kDa)*					
sapi3_1/ear	185	20.2	1 (*ear*)	*sel*		Y	Y
seb	266	31.4		*sec*	Enterotoxin	Y	Y
sapi3_3	183	20.7	2	5	(Terminase)	Y	
sapi3_4	113	13.4	3	6		Y	
sapi3_5	175	20.6	4	7		Y	
sapi3_6	72	8.2	5	8		Y	Y
sapi3_7	192	22.8	6	9			
sapi3_8	113	13.4		10		Y	
sapi3_9	213	24.5	7	11		Y	
sapi3_10	94	10.8	8				
sapi3_11	120	13.9		12		Y	
sapi3_12	474	55.0	11		(*vapE*)	Y	
sapi3_13	289	33.3	13/12	15	(Replication protein)	Y	
sapi3_14	106	12.7		16		Y	Y
sapi3_15	69	8.3				Y	Y
sapi3_16	55	6.0				Y	Y
sapi3_17	90	10.4	14		(ORF 37 *S. aureus* phage ϕPVL)	Y	
sapi3_18	66	7.6			(Cro-like repressor)	Y	Y
sapi3_19	110	12.7			(cI-like repressor)	Y	Y
sapi3_20	152	18.1			(ORF 153 of ϕSLT)	Y	Y
sapi3_21	143	16.2			(ABC transporter)	Y	Y
seq	242	28.2	*seq′*		Enterotoxin	Y	Y
sek	242	27.7	*sek*		Enterotoxin	Y	Y
sapi3_int	407	47.6	*int*	*int*	Integrase	Y	Y

[a] Determined using the Compute pI/Mw tool at ca.expasy.org/tools/pi_tool.html (51).
[b] Corresponding ORFs were determined by relative position within the islands and/or similarity of the predicted gene products.

SaPI3 contains this 100-bp stretch as well as an additional 67 bp not present in either SaPI1 or SaPIbov (nucleotides 6063–6230). The second core region includes nucleotides 9380–10,284 of SaPI1, 9591–10,494 of SaPI3, and 11,419–12,429 of SaPIbov. All three islands contain *int* genes adjacent to the *attR* sites, and SaPI3 contains an enterotoxin gene (*seb*) in the same position as *tstH* in SaPI1 and SaPIbov. SaPI1 and SaPI3 appear to be even more closely related. In addition to the shared elements among all three islands and the identical attachment sites in SaPI1 and SaPI3, these two islands are highly homologous (>93% identity) at the right end (nucleotides 10,285–11,046 in SaPI1 and 10,494–11,254 in SaPI3; nucleotides 12,248–15,250 in SaPI1 and 12,891–15,953 in SaPI3) (Fig. 1). The *ear* (*sapi1_1* and *sapi3_1*) and *sek* genes are also in the same relative positions to each other and to the *att* sites of SaPI1 and SaPI3. Although the function of *ear* is unknown, several properties of the gene suggest that it may have an important function in the life cycle of *S. aureus*. Its position and predicted product are conserved (~75% identity at the amino acid level) among SaPI1, SaPI3, and SaPI4 (data not shown), it has the identical signal sequence as TSST-1, and it is secreted in abundant quantities by *S. aureus* RN4282. In addition to homology between SaPI1, SaPI3, and SaPIbov, the region encoding *sapi3_3-9* and *sapi3_15* shares extensive homology (>95% at the nucleotide level) to a matching region in SaPIn1/SaPIm1, whereas *sapi3_10-14* are 87% or more similar to regions presumably of phage origin in strain Mu50.

A locus of ~900 bp adjacent to the *attL* site of SaPI3 was 95% identical to a phage 80α sequence adjacent to the putative phage amidase gene. However, this region of SaPI3 apparently does not encode for any protein. A ~1.6-kb region SaPI3 with significant variance as compared with SaPI1 is found spanning the region between *sapi3_17* and *seq*. Within this region is a stretch of ~500 nucleotides that is nearly identical to a sequence from the recently identified ϕSLT, a temperate *S. aureus* phage encoding the Panton-Valentine leukocidin. Immediately upstream of *int* in both SaPI1 and SaPI3 is a 46-bp sequence conserved among staphylococcal phages ϕ11, ϕ13, ϕ42, and L54a (18–20). This sequence is the binding site for two phage ϕ11 proteins that regulate *int* expression, RinA and RinB (21).

In addition to the integrase and the noncoding regions with homology to phage DNA, several of the genes contained on SaPI3 suggest a mobile element of phage origin, perhaps a conglomeration of several phage elements (Table I). The terminase potentially encoded by *sapi3_3* is homologous to the small subunit of identified terminases in the bacteriophages ρ15 (22) (52% similarity over 162 amino acids) and PBSX of *Bacillus subtilis* (23) (50% similarity over 102 amino acids). *Sapi3_12* potentially encodes a product with high homology (54% similarity over 333 amino acids) to the virulence-associated protein (VapE) of *Dichelobacter nodosus*, a sheep pathogen (24). (*Sapi1_11* is also a VapE homolog 9.) This has led to the supposition that this gene was acquired by *S. aureus* through horizontal transfer from *D. nodosus* during co-colonization or infection of sheep (5) because the *D. nodosus vap* genes appear to reside on an integrated bacteriophage (25, 26). The predicted product of *sapi3_18* is 75% similar over 58 amino acids to a putative *cro*-like repressor of *Streptococcus thermophilus* bacteriophage Sfi21 (27, 28). Even stronger homology is seen between the predicted product of *sapi3_19* and the *cI*-like repressor of *S. thermophilus* phage Sfi21 (27, 28) (84% similarity over 67 amino acids) and *Lactobacillus casei* phage A2 repressor (29) (80% similarity over 76 amino acids). *Sapi3_20* potentially encodes a glycoprotein similar to one found in bacteriophage A118 (30) (54% similarity over 123 amino acids) and is strongly similar to ORF 135 from the recently described *S. aureus* temperate phage ϕSLT carrying Panton-Valentine leukocidin (31) (85% similarity over 153 amino acids).

Expression Analysis of SaPI3—Studies of pathogenicity islands in other bacterial species have demonstrated that genes contained on the islands act in regulation of virulence factors carried by the island (reviewed in Ref. 7). SaPI3 contains several ORFs with no identifiable function, and it was not known whether or not they were expressed, and, if so, whether they might act in regulation of the SaPI3 enterotoxins, *seb*, *sek*, and

TABLE II
Primers used to PCR-amplify SaPI3 microarray probes[a]

Gene name	Forward primer (5′-3′)[b]	Reverse primer (5′-3′)[c]
sapi3_1/ear	GTAGTATTACGGGTACAGC	gatcggatccTTATTTAGTTATAGTTATTTTTG
seb	GAAGTATTACTGTTCGGG	gatcggatccTCACTTTTTCTTTGTCGTAA
sapi3_3	ATGAAACAGAAACGAAAG	gatcggatccTCAATCCTCTAACGGAATA
sapi3_4	GATCCCATGGCTATGGATAAAAAGCAAATAAAA	gatcggatccTCTATCACCTCCACAATTTT
sapi3_5	CCATGGCTATGAGTGCGTCGACCAC	gatcggatccCTAACTGTTTTCAATTGCT
sapi3_6	CCATGGCTATGGAAACAAAATGCGAG	gatcggatccCTATTTAATAATCCTGTTTTG
sapi3_7	AGACAATTGAATTATTAAAACA	gatcggatccTTAACGTTTTAAAAACAACTTG
sapi3_8	CCATGGCTATGCAAAGTATCGCAGAA	gatcggatccTTACACTTTAATTCGTTTCA
sapi3_9	ATGATGGCCAGTCAATTG	gatcggatccTTATTTACTCAAGCTATAGT
sapi3_10	CCATGGCTATGAATGATAAAGAGAAAATTT	gatcggatccTGTGTTATCTCCTTTTGAG
sapi3_11	CCATGGCTATGAATAAAAATCAATTAAAGTC	gatcggatccTTAGACACACTCCGTTTC
sapi3_12	TGGTATAGAACGTTAGAAG	gatcggatccTTAAACTTTAAGATTCTTATAA
sapi3_13	GTTGATATGCTCGAACAG	gatcggatccCTAACTGTTTTTACTATCTT
sapi3_14	ccatggctATGAATTGGGAAATTAAAGA	gatcggatccTTACTCTGCTACTTTTATTG
sapi3_15	ccatggctATGAATAATGAACAAAAAGAA	gatcggatccTTATTCTCCTTTCTCTAATT
sapi3_16	ccatggctATGGATGTAGGAGGCAAG	gatcggatccTTATTCATTTTCTAATCCTCC
sapi3_17	ccatggctATGTCTAGAACAAAATTGCAT	gatcggatccCTACATCCATTTTTTATGAC
sapi3_18	ccatggctATGGCTCACAAAATGGGG	gatcggatccCTAGACATTTAATTCACCTC
sapi3_19	ccatggctATGAGAACTAATGATGAAAT	gatcggatccTTACTGATTTCGATGTGCTT
sapi3_20	ACGAAGAACTTGCCCAC	gatcggatccTTATATTTCTTTATATTTAAAAAC
sapi3_21	AAGATAATAAATTAAATGTTATC	gatcggatccTCACTTTAATGTTTTATCCAT
seq	AAGAGGTAACTGCTCAAG	gatcggatccTTATTCAGTCTTCTCATATG
sek	ACCGCTCAAGAGATTGAT	gatcggatccTTATATCGTTTCTTTATAAGAA
sapi3_int	ATTCAAAATGGAATGATGGA	gatcggatccTTATCCAAGTTTTACCTGT

[a] The forward direction for each primer pair was determined by the orientation of the corresponding gene.
[b] Some forward primers incorporate a *Nco*I restriction site (lowercase characters).
[c] Reverse primers incorporate a *Bam*H1 restriction site (lowercase characters).

seq, or only in the maintenance and transfer of the island itself. We thus used DNA microarrays to examine the expression of all SaPI3 ORFs potentially encoding products over 50 amino acids in size to determine whether or not they were detectably expressed and whether their expression profiles were similar to those of the SaPI3-associated enterotoxins. DNA primers used to PCR-amplify probes for each ORF are listed in Table II.

Expression data for SaPI3 genes in various growth conditions are summarized in Table III. Because staphylococcal exotoxins are generally growth phase-regulated, we first examined the effect of growth stage on the expression of SaPI3 genes. In all, the expression of 13 of 24 ORFs was detected during growth of MN NJ in TH broth. The expression of five of these genes was significantly ($p < 0.05$, Student's t test) altered by postexponential growth as compared with exponential phase growth, whereas the expression of four genes was affected by stationary phase growth. As expected, the expression of *seb*, known to be a postexponential and stationary phase-produced exotoxin, was increased by 8.2-fold in stationary phase as compared with the exponential phase of growth. Interestingly, the expression of *sek* and *seq* was unaffected by growth phase, suggesting that these are constitutively produced exotoxins in laboratory media.

Previous studies have demonstrated repression of another staphylococcal toxin associated with TSS, TSST-1, by anaerobic conditions and enhancement of toxin production in aerobic conditions, particularly in the presence of elevated carbon dioxide (15, 32–35). Our laboratory has also determined that the effect of oxygen on *tstH* expression occurs primarily at the transcriptional level and is independent of cell density and pH (15).[3] To examine the effect of oxygen concentration on the expression of SaPI3 genes, the cultures were exposed to microaerobic (1% O_2 (v/v)) and aerobic (21% O_2 (v/v)) growth conditions in 1-ml cultures of TH broth. There was no significant difference in growth parameters between the microaerobic and aerobic cultures because both cultures entered their respective growth phases (exponential, postexponential, and stationary) of interest at similar times. Also, differences in cell densities between the microaerobic and aerobic cultures were 2-fold or less in any given growth phase. In all, the expression of 18 of the 24 SaPI3 ORFs was detected. The expression of four, four, and three genes was significantly affected by growth in microaerobic conditions in the exponential, postexponential, and stationary phases of growth, respectively. The expression of *seb* was somewhat repressed in the exponential and postexponential phases of growth in aerobic conditions and was slightly up-regulated during stationary phase growth in aerobic conditions. The expression of no gene, however, was affected by more than approximately 2-fold.

To approximate *in vivo* conditions and determine those genes whose expression might be artificially enhanced by growth in rich laboratory medium, the expression of SaPI3 genes was examined during growth of MN NJ in TH broth *versus* rabbit serum. Effect of growth in serum was examined in the exponential, postexponential, and stationary phases of growth. Cells were quantified by both absorbance and dry weight. No difference in the time at which TH or serum cultures entered their respective growth phases of interest was observed. $A_{600\ nm}$ values for *S. aureus* grown in TH medium were 0.83, 1.69, and 2.16 for the exponential, postexponential, and stationary phases of growth, respectively. $A_{600\ nm}$ values for *S. aureus* grown in rabbit serum were 0.33, 0.97, and 1.27 for the same growth phases. As measured by dry weight, cell mass at 8 h was 8.0 mg/ml in TH medium and 8.3 mg/ml in rabbit serum. Quantification of *S. aureus* growth in serum *in vitro* is problematic because cells clump tightly and are resistant to dispersal, even by ultrasonication, thus preventing accurate counts by plating or absorbance. Alternatively, dry weight analysis is hampered by the tendency of *S. aureus* to bind serum components and thus add "artificial" cell mass. Thus, an effect of cell density on SaPI3 gene expression cannot be ruled out in these serum cultures. Taken as a whole, however, the data reflect the fact that growth in standard rich laboratory media may well enhance or repress expression of virulence-associated genes in a manner inconsistent with what occurs *in vivo*. In all, the expression of 19 of 24 SaPI3 ORFs was detected in these experiments. The expression of five, six, and three genes were affected by growth in serum as compared with TH

[3] J. M. Yarwood and P. M. Schlievert, unpublished observations.

TABLE III
Expression data for SaPI3 ORFs[a]

Bold and italic fonts are used to indicate those genes whose expression was significantly increased or decreased, respectively, in the test experimental condition (bold, italic font) as compared to the reference experimental condition (underline). Blank spaces indicated those ORFs whose expression was not detected under the growth conditions examined.

Gene	Growth phase		***Microaerobic***/aerobic			***Serum***/laboratory medium			***In vivo***/inoculum		
	P-exp/Exp	***Stat***/Exp	Exp	P-exp	Stat	Exp	P-exp	Stat	N-Imm 2 h	Imm-2 h	Imm-8 h
sapi3_1/ear	**1.7 ± 0.2**[b]	1.1 ± 0.2	*−1.9 ± 0.1*[b]	**1.7 ± 0.1**[b]	**1.5 ± 0.1**[b]	−1.4 ± 0.3	*−3.3 ± 1.1*[b]	*−2.0 ± 0.2*[b]	**3.4 ± 0.5**[b]	**1.6 ± 0.1**[c]	*−1.4 ± 0.1*[b]
seb	**6.4 ± 1.4**[b]	**8.2 ± 1.1**[b]	*−2.1 ± 0.2*[b]	*−1.6 ± 0.2*[b]	**2.2 ± 0.5**[b]	1.3 ± 0.3	*−2.1 ± 0.1*[b]	*−1.6 ± 0.2*[b]	**3.6 ± 0.2**[b]	**1.8 ± 0.2**[b]	−1.3 ± 0.1
sapi3_3					1.0 ± 0.1	1.4 ± 0.1	1.3 ± 0.4	1.3 ± 0.5			
sapi3_4			1.3 ± 0.3		−1.3 ± 0.2	1.4 ± 0.1	−1.1 ± 0.1				
sapi3_5				1.2 ± 0.2	−1.1 ± <0.1	1.7 ± 0.3	−1.4 ± 0.3				
sapi3_6		2.0 ± 0.5	1.2 ± 0.2	1.2 ± 0.1	1.1 ± 0.1	1.0 ± 0.2	−1.1 ± 0.2	1.2 ± 0.1	−1.5 ± 0.1	1.1 ± 0.2	**1.3 ± 0.1**[b]
sapi3_7											
sapi3_8		**2.5 ± 0.5**[b]	**1.7 ± 0.3**	**1.2 ± 0.1**[b]	1.1 ± <0.1	1.5 ± 0.3					
sapi3_9					−1.1 ± <0.1						
sapi3_10											
sapi3_11	−1.2 ± 0.1						−1.2 ± 0.3				
sapi3_12	*−2.2 ± 0.3*[b]				−1.1 ± 0.1						
sapi3_13						2.2 ± 0.5					
sapi3_14				1.2 ± 0.1	**1.7 ± 0.3**[b]	1.3 ± 0.3		−1.1 ± 0.1	*−1.5 ± 0.1*[b]		−1.4 ± 0.1
sapi3_15		1.3 ± 0.2			1.3 ± 0.2						−1.2 ± 0.2
sapi3_16					1.1 ± 0.1		−1.3 ± 0.4				−1.2 ± 0.1
sapi3_17							*−1.5 ± 0.2*[b]				
sapi3_18					1.1 ± 0.1	1.2 ± 0.2	1.2 ± 0.2	**1.6 ± 0.2**[b]		−1.3 ± 0.1	−1.3 ± 0.2
sapi3_19	**1.8 ± 0.2**[b]	*−2.4 ± 0.8*[b]	1.6 ± 0.6	−1.5 ± 0.1	1.0 ± 0.1	1.4 ± 0.2	−1.5 ± 0.2	−1.1 ± 0.1	1.0 ± 0.2	−1.6 ± 0.2	
sapi3_20	**1.5 ± 0.2**[b]	*−2.7 ± 0.7*[b]	−1.1 ± 0.1	−1.3 ± 0.1	−1.2 ± 0.1	**2.0 ± 0.2**[b]		−1.4 ± 0.1	−1.4 ± 0.2	−1.3 ± 0.2	*−2.7 ± 0.7*[b]
sapi3_21	−1.1 ± 0.2	−1.5 ± 0.4	1.2 ± 0.1		−1.2 ± 0.1	**1.6 ± 0.1**[b]	*−2.2 ± 0.4*[b]	1.1 ± 0.1	−1.4 ± 0.2	−1.4 ± 0.2	*−2.3 ± 0.5*[b]
seq	−1.3 ± 0.3		1.2 ± 0.2		1.1 ± 0.1	**2.3 ± 0.6**[b]	*−1.7 ± 0.4*[b]	1.4 ± 0.1	−1.2 ± <0.1	−1.4 ± 0.2	*−2.1 ± 0.5*[b]
sek	1.1 ± 0.1	−1.3 ± 0.2	1.2 ± 0.2	*−1.6 ± 0.1*[b]	1.1 ± 0.1	**1.6 ± <0.1**[b]	*−1.7 ± 0.1*[b]	−1.4 ± 0.2	*−1.5 ± 0.2*[b]	*−1.6 ± 0.2*[b]	*−3.7 ± 0.5*[b]
sapi3_int		−1.2 ± 0.1	1.6 ± 0.2[b]	−1.3 ± 0.3	1.2 ± 0.1	**2.4 ± 0.4**[b]	−1.1 ± 0.2	1.9 ± 0.1		*−1.8 ± 0.1*[b]	*−2.3 ± 0.2*[b]

[a] Exp, exponential phase growth; P-exp, postexponential phase growth; Stat, stationary phase growth; N-Imm, nonimmune; Imm, immune.
[b] $p < 0.05$, Student's t test.
[c] $p < 0.01$, Student's t test.

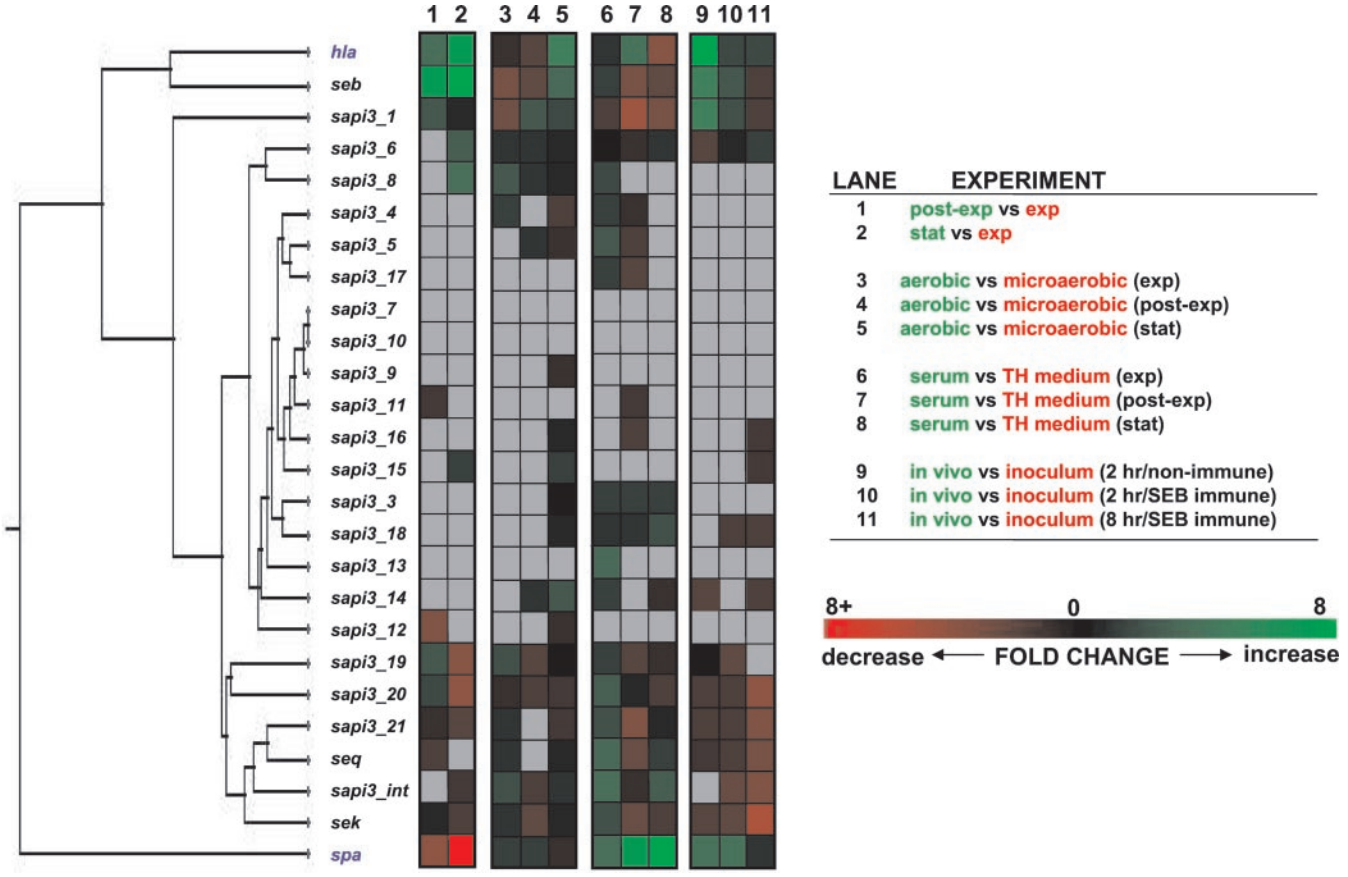

FIG. 2. **Expression profiles and hierarchical clustering of the genes for α-hemolysin (*hla*) and protein A (*spa*) and the 24 SaPI3 ORFs potentially encoding products of more than 50 amino acids.** *Red* and *green* represent fold decrease and increase, respectively, in gene expression in response to later growth phase (*lanes 1* and *2*), growth in microaerobic *versus* aerobic conditions (*lanes 3–5*), growth in rabbit serum *versus* laboratory medium (*lanes 6–8*), and incubation *in vivo* as compared with the inoculum (*lanes 9–11*). *Gray shading* indicates those ORFs whose expression was not detected in the corresponding growth condition. Clustering based on similarity of expression profiles and visualization were performed using the software program Spotfire DecisionSite 6.1 (www.spotfire.com). Similarities between expression profiles of individual genes in all 11 experimental conditions were determined using the Euclidean distance method. *exp*, exponential; *post-exp*, postexponential; *stat*, stationary.

medium in the exponential, postexponential, and stationary phases of growth, respectively, although no gene was affected by more than 3.3-fold. In general, expression of genes on either end of the island was more consistently detected than *sapi3_7-17*.

Finally, we examined the expression of SaPI3 genes during incubation of MN NJ *in vivo* in a rabbit model of TSS using subcutaneous hollow polyethylene infection chambers (17, 36). The expression of SaPI3 genes *in vivo* was compared with that in the inoculum, which was harvested from cultures in the late exponential phase of growth and used to infect both nonimmune rabbits and rabbits immunized with SEB. All animals developed symptoms consistent with TSS, including hypotension, respiratory distress, and obvious discomfort in the vicinity of the infection chamber. The SEB-immunized rabbits experienced a delayed onset of symptoms and were euthanized 22 h after inoculation, whereas the nonimmune rabbits died several hours after inoculation. Cell densities recovered from the infection chambers did not vary more than 0.22 log units from the cell density of the inoculum before concentration (6.7×10^8 colony-forming units/ml), thus we were able to effectively eliminate the potential effects of cell density and growth phase on gene expression *in vivo*. Using this subcutaneous infection model, we were previously able to demonstrate up to 18-fold changes in virulence-associated gene expression between cells in the inoculum and cells *in vivo* (36). The expression of 13 of 24 SaPI3 ORFs was detected *in vivo*. *Sapi3_1* (*ear*) and *seb* were significantly up-regulated during incubation *in vivo* in the nonimmune rabbits as compared with the inoculum and were affected very little by incubation in the immune animals. The expression of *sek* and *seq* was only significantly affected by incubation of MN NJ in the immune animals at 8 h, in which case they were repressed, together with *sapi3_20*, *sapi3_21*, and *sapi3_24*.

To determine whether any of the genes contained on SaPI3 might act in the regulation of or be co-regulated with the enterotoxins carried by the island, the genes were clustered according to the similarity of their expression profiles (Fig. 2). The expression profiles of the secreted toxin α-hemolysin (*hla*) and the surface molecule protein A (*spa*) in these experimental conditions are provided as representative virulence factors. To our knowledge, the presence and expression of *hla* and *spa* are not affected by the presence of SaPI3 or any of the other described pathogenicity islands. Three primary clusters of SaPI3 genes were observed. The expression profile of *seb* matched most closely with that of *hla* and not any of the other SaPI3 genes. In contrast, the expression of *sek* and *seq* was found to be most similar to that of the surrounding genes (*sapi3_19-21* and *sapi3_24*). This clustering of *sapi3_19-24* in their expression profiles is suggestive because the direction of transcription of these genes is opposite that of nearly all of the remaining genes (*sapi3_2-18*) on SaPI3. Finally, those genes centrally located in the island (*sapi3_3-18*, with the exception of *sapi3_6*), whose expression was detected only in a limited number of growth conditions, form a third group with variable expression patterns. Although it was consistently detected,

sapi3_1 (*ear*) did not group with any of the other genes examined.

Although the resolution of these array experiments did not allow a conclusive determination that any of the SaPI3 genes are co-transcribed, the proximity of several ORFs to one another and the clustering of their expression profiles are suggestive. For instance, *sapi3_4* and *sapi3_5,* which are separated by only 2 nucleotides, cluster together, as do *sapi3_19* and *sapi3_20*, separated by 11 nucleotides. Interestingly, *sapi3_21* and *seq*, separated by 11 nucleotides, also have similar expression profiles. Genes *sapi3_9-11* form a set of overlapping or nearly overlapping ORFs, but their expression was not detected in a sufficient number of conditions to make a conclusion regarding their possible co-transcription. Genes *sapi3_14-18* also form a set of overlapping or nearly overlapping reading frames; however, only *sapi3_15* and *sapi3_16*, which overlap by 13 nucleotides, cluster in their expression profiles. The presence of co-transcribed genes would not be unexpected if the pathogenicity island was indeed of phage origin. Co-transcribed genes are common in phage genomes because numerous products, such as structural elements, are encoded by polycistronic transcripts to achieve coordinate temporal regulation at critical points within the phage lytic cycle. The presence of several stem-loop structures that may serve as rho-independent transcription terminators is indicated in Fig. 1. The placement of these structures is consistent with those genes that are known or likely to be independently transcribed, such as *seb*, *sapi3_1* (*ear*), *sek*, and *int*, as well as with those genes that may well be co-transcribed, such as *sapi3_14-18*.

In all, we were able to detect the expression of 22 of the 24 SaPI3 ORFs by microarray analysis (Table I). 13 ORFs were detectably expressed *in vivo* and *in vitro*, 9 ORFs were detectably expressed *in vitro* only, and 2 ORFs were not detectably expressed in any growth condition examined. The expression of genes toward either end of SaPI3 was more consistently detected than that of genes contained in the central region of the island. Consistent with a role for these islands in dissemination of staphylococcal enterotoxins, the expression of the toxins *seb*, *sek*, and *seq* was consistently detected under multiple growth conditions.

DISCUSSION

In the work presented here, we have described a novel pathogenicity island, SaPI3, in *S. aureus* strains COL and MN NJ. We propose that the newly identified island meets the consensus definitions of a pathogenicity island (7), including the presence of demonstrated virulence genes on the island (*seb*, *sek*, and *seq*), the lack of the island in closely related strains, the occupation of a relatively large genomic region (~16 kb), a lower GC content than the overall *S. aureus* genome (31.4% *versus* 32.8%), the presence of flanking direct repeats, and the presence of mobility factors (integrase). Many of the genes contained on the island are homologous to genes contained on described bacteriophages, suggesting that this SaPI, like others described previously, is of bacteriophage origin. Furthermore, we have detected the expression of 22 of the 24 genes contained on the island, suggesting that many of them may be active in the maintenance of the island or in regulation of the associated enterotoxins.

Several explanations for the origin of the staphylococcal pathogenicity islands are possible. A large number of pathogenicity islands may circulate through the combined gene pool of numerous bacterial species with which staphylococci is transiently associated, in or on its human and animal hosts, and from which staphylococci acquires these elements through horizontal transfer. However, this hypothesis seems unlikely because superantigen production has been identified in a very limited number of bacterial species, suggesting that these elements are not in general circulation. The possibility cannot be eliminated completely, however, until the genomes of currently unidentified bacterial associates of staphylococci are characterized. A second hypothesis suggests that exotoxins are carried primarily by other genetic elements, such as plasmids and phage, and have integrated into sites of high recombination frequency in one more ancestral strains. These recombination events might be mediated by homologous genes in the accessory genetic elements and the pathogenicity island, giving rise to the exotoxin-carrying islands. However, this hypothesis does not account for the origin of superantigen genes on transmissible elements, the apparently exclusive presence of certain superantigen genes on only pathogenicity islands, and the likely role for the demonstrably mobile pathogenicity islands in dissemination and evolution of superantigens. A third hypothesis, which we favor, combines certain elements of the first two, with the addition of the assumption of a common ancestral genetic element to SaPI1–4, SaPIbov, and SaPIm1/SaPIn1 and the continual generation of unique toxins and islands through modular recombination events.

Several features of the enterotoxin-encoding pathogenicity islands identified thus far suggest a common ancestral genetic element and that these islands have arisen in part through specialized transduction and recombination events. The overall layout of the islands, even specific genes, is similar, with exotoxins encoded on either end of the islands (the location of *tstH* is nearly identical in SaPIn1/SaPIm1, SaPI1, and SaPIbov), and the exotoxin genes on the right end of the islands consistently located upstream of the integrase genes. Toward the center of the islands lie multiple genes of apparently phage origin. Furthermore, the region in SaPI3 between nucleotides 3113 and 6929 and nucleotides 9591 and 10,494 is nearly identical to regions in all of these islands. Thus, we propose a generalized model for the origin and evolution of staphylococcal pathogenicity islands (Fig. 3). The presence of the exotoxin genes on the left ends of the islands can be explained by a specialized transduction event involving an ancestral, exotoxin-free bacteriophage in either *Staphylococcus* or a bacterial species with which *Staphylococcus* has subsequently engaged in horizontal gene transfer. In this model, a bacteriophage element excised from the chromosome adjacent to a superantigen locus and through a mis-recombination event gained the enterotoxin while either simultaneously or subsequently losing a segment of phage DNA necessary for complete phage function (such as *xis*). It thus became dependent upon wild-type helper phage for excision, packaging, and/or mobilization, as observed with SaPI1. Indeed, the expression of *seb*, which is known to be regulated by the staphylococcal accessory gene regulator (*agr*) *in vitro*, did not correlate with the expression of any other gene contained on the island but rather correlated with the *agr*-regulated gene *hla*, which is not associated with SaPI3. This supports the hypothesis that these exotoxins were incorporated into an existing phage.

The presence of exotoxins on the right end of the island, inside the integrase, may instead be due to addition, deletion, or mutual crossover of large pathogenicity island fragments or "modules." The hypothesis of modular recombination events perhaps mediated by key regions of sequence homology is supported by both the presence of long stretches of homology between distinct regions of the islands and the expression data described in this work. Extensive regions of very high similarity (>95% identity), such as that between SaPI3 and SaPI1, are abruptly interrupted by regions with very little homology between the islands. Some of these regions appear to be derived from phages such as Panton-Valentine leukocidin-converting

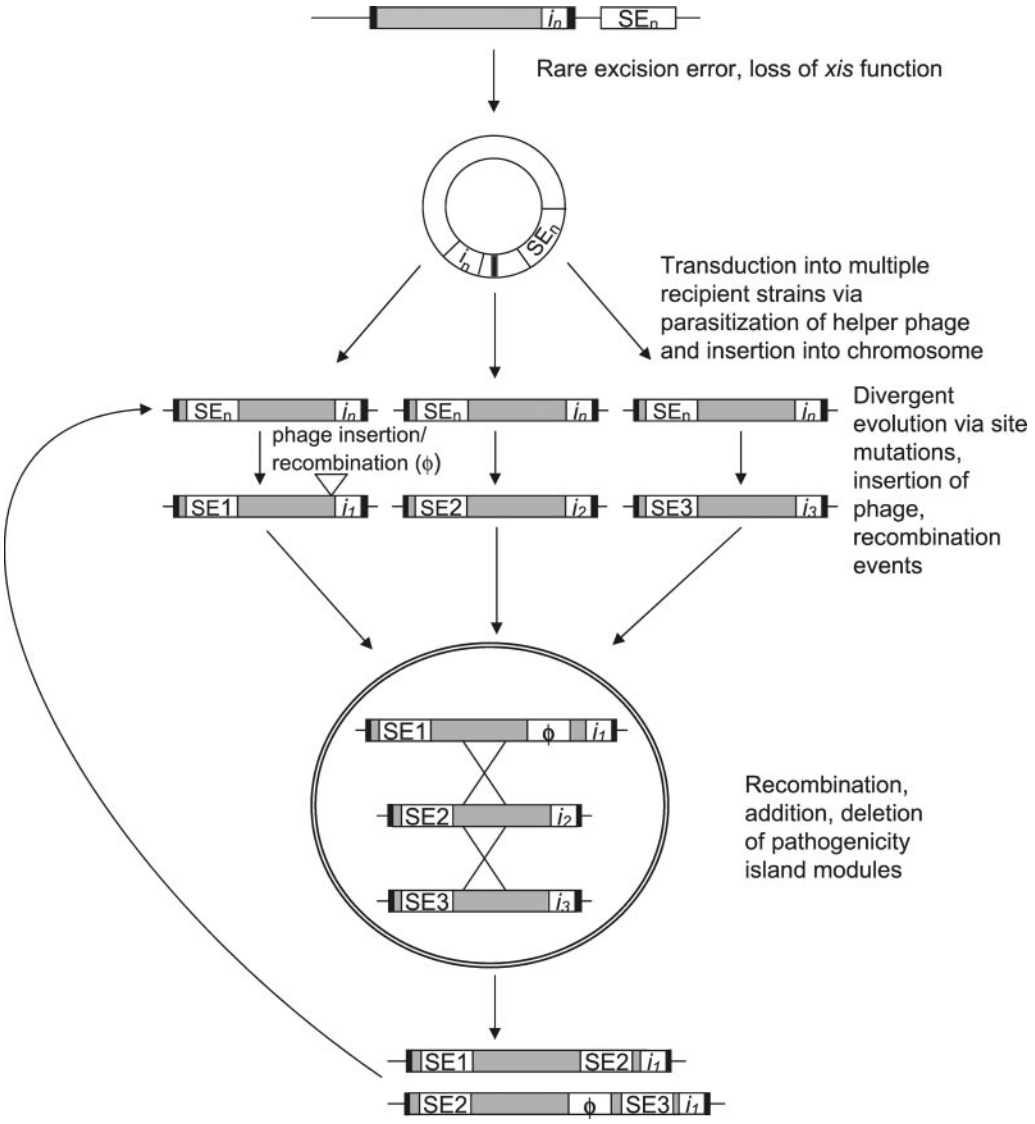

FIG. 3. **A generalized model for the origination and evolution of staphylococcal pathogenicity islands.** A rare excision error by a fully functional bacteriophage leads to the addition of an ancestral exotoxin (SE_n) gene to the phage genome and loss of critical phage functions. The defective phage, now dependent on helper phage for transfer into recipient strains, undergoes divergent evolution in multiple staphylococcal lineages. Distinct enterotoxins arise, and integration sites may be altered through evolution of the integrase genes. As distinct pathogenicity islands then insert into identical host strains, addition, deletion, and recombination of large segments or modules of the pathogenicity islands occur. The cycle is repeated in various permutations as the resulting pathogenicity islands then undergo further divergent evolution in subsequent staphylococcal lineages. *int*, integrase gene; *SE*, enterotoxin gene; *ϕ*, phage element.

phage ϕSLT. Furthermore, SaPI1 and SaPI3 appear to have lost the *sel/sek/sec3* module present on SaPIn1/SaPIm1 and SaPIbov, or, alternatively, SaPIn1/SaPIm1 and SaPIbov have gained these elements after diverging from a common ancestor. In addition, we found that we could consistently detect expression of genes on either end of the island, particularly *sapi3_19-24*. However, the "core" bacteriophage genes were much more difficult to detect and may only be strongly expressed upon mobilization of the island. Although most of these central genes have no definite homolog, they may well encode proteins involved in the structural components of the original phage. In phage λ, for instance, genes encoding head and tail components are located together at the end of the prophage opposite the *int* and *xis* genes (37). Thus, these genes may only be expressed when the pathogenicity island is mobilized. Alternatively, the promoter regions for these genes may be defective or even absent, preventing the expression of these genes at any time and requiring the structural components of a helper phage for transfer. The observation, then, that *sapi3_19-24* are transcribed opposite to *sapi3_3-18* and cluster in their expression profiles implies co-regulation of these genes and the addition of this enterotoxin module to the already existing pathogenicity island. Interestingly, the DNA sequence containing *sapi3_3-17* is highly conserved between SaPI1 and SaPI3 and, to a lesser extent, SaPIbov. It is not yet clear whether the size and type of these modules might be constrained by recombination events that require conserved stretches of DNA sequences in the phage genomes or are more or less randomly generated. The possibility also cannot be excluded that staphylococcal pathogenicity islands have arisen in part as a result of horizontal transfer from another species. Indeed, streptococcal pyrogenic exotoxin A, a phage-encoded superantigen, appears to be more closely related to staphylococcal enterotoxins SEB, SEC3, and SEG, than other streptococcal superantigens (4).

We also propose that the promiscuity of these islands and their tendency to undergo recombination events underlie the evolutionary divergence of staphylococcal superantigens. The presence of an ancestral island in multiple staphylococcal lineages undergoing separate genetic events (*e.g.* point mutations, genetic rearrangements, and insertion of other mobile genetic

elements) and evolutionary pressures in various communities and hosts would lead to a great diversity of exotoxin genes as well as pathogenicity islands (Fig. 3). Further recombination events upon mixing of these various staphylococcal lineages would lead to even greater diversity among the pathogenicity islands and perhaps other mobile genetic elements as well. However, the evolutionary fitness that superantigenic toxins confer to the recipient strains is not yet completely understood. It is clear that these superantigens have significant immunodeletory effects, including inducing anergy and deletion of a large population of T cells (38–40) as well as preventing the development of antibody to the toxin itself (41–44). There is also evidence that superantigens destroy endothelial cells (45) and that they may exclude neutrophils from infection foci (46–48). Thus, superantigens likely confer immunological protection to their host strains and perhaps prevent clearance of those strains from the human host.

The demonstrated mobility of SaPI1 and the presence of apparently functional integrases in all of these islands (5, 9, 10) provide further support for the hypothesis that they remain a significant part of the evolutionary scheme of *S. aureus* and will likely give rise to new enterotoxins and pathogenicity islands. The mobility of SaPIs has led to speculation that these islands represent relatively "young," recently acquired genetic elements, as opposed to islands that likely entered their host organisms millions of years ago and have become relatively immobile, perhaps even part of the host core genome (7). Alternatively, the ability of staphylococci and their associated pathogenicity islands to evolve may be constrained by the evolution of their human hosts. Thus, the conversion of any particular SaPI into a stable element may be restricted by the need to constantly respond to the adaptive human immune response. Constant generation of new superantigens and pathogenicity islands might enable *S. aureus* to colonize and infect human populations that may already have acquired immunity to ancestral exotoxins. Indeed, lack of antibody to exotoxin is a key risk factor for the development of staphylococcal TSS (41, 49, 50).

As additional *S. aureus* genome sequences and SaPIs become available and are associated with various staphylococcal lineages and infection types, it may become easier to identify a common ancestor of staphylococcal pathogenicity islands and superantigens. Regardless, continued research into the mechanisms of superantigen and SaPI evolution and acquisition will allow more accurate epidemiological investigations as well as a greater understanding of which potential therapeutic interventions might be viable, particularly in the development of toxoid vaccines.

Acknowledgments—We thank Barbara May for assistance with reverse transcription-PCR and Peter Southern for helpful discussions regarding SaPI evolutionary models. Preliminary sequence data were obtained from The Institute for Genomic Research website (www.tigr.org) with support provided by the National Institute of Allergy and Infectious Diseases and Merck Genome Research Institute.

REFERENCES

1. Tenover, F. C., and Gaynes, R. P. (2000) in *Gram-positive Pathogens* (Fischetti, V. A., Novick, R. P., Ferretti, J. J., Portnoy, D. A., and Rood, J. I., eds), pp. 414–421, ASM Press, Washington, D. C.
2. Marrack, P., and Kappler, J. (1990) *Science* **248,** 705–711
3. Hovde, C. J., Marr, J. C., Hoffmann, M. L., Hackett, S. P., Chi, Y. I., Crum, K. K., Stevens, D. L., Stauffacher, C. V., and Bohach, G. A. (1994) *Mol. Microbiol.* **13,** 897–909
4. McCormick, J. K., Yarwood, J. M., and Schlievert, P. M. (2001) *Annu. Rev. Microbiol.* **55,** 77–104
5. Novick, R. P., Schlievert, P., and Ruzin, A. (2001) *Microbes Infect.* **3,** 585–594
6. Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H., and Hiramatsu, K. (2001) *Lancet* **357,** 1225–1240
7. Hacker, J., and Kaper, J. B. (2000) *Annu. Rev. Microbiol.* **54,** 641–679
8. Hentschel, U., and Hacker, J. (2001) *Microbes Infect.* **3,** 545–548
9. Lindsay, J. A., Ruzin, A., Ross, H. F., Kurepina, N., and Novick, R. P. (1998) *Mol. Microbiol.* **29,** 527–543
10. Ruzin, A., Lindsay, J., and Novick, R. P. (2001) *Mol. Microbiol.* **41,** 365–377
11. Moore, P. C., and Lindsay, J. A. (2001) *J. Clin. Microbiol.* **39,** 2760–2767
12. Orwin, P. M., Leung, D. Y., Donahue, H. L., Novick, R. P., and Schlievert, P. M. (2001) *Infect. Immun.* **69,** 360–366
13. Schlievert, P. M., and Blomster, D. A. (1983) *J. Infect. Dis.* **147,** 236–242
14. Mishell, B. B., and Mishell, R. I. (1980) in *Selected Methods in Cellular Immunology* (Mishell, B. B., and Shiigi, S. M., eds), pp. 30–37, W. H. Freeman and Company, San Francisco
15. Yarwood, J. M., and Schlievert, P. M. (2000) *J. Clin. Microbiol.* **38,** 1797–1803
16. Yarwood, J. M., McCormick, J. K., and Schlievert, P. M. (2001) *J. Bacteriol.* **183,** 1113–1123
17. Scott, D. F., Kling, J. M., Kirkland, J. J., and Best, G. K. (1983) *Infect. Immun.* **39,** 383–387
18. Carroll, D., Kehoe, M. A., Cavanagh, D., and Coleman, D. C. (1995) *Mol. Microbiol.* **16,** 877–893
19. Ye, Z. H., Buranen, S. L., and Lee, C. Y. (1990) *J. Bacteriol.* **172,** 2568–2575
20. Ye, Z. H., and Lee, C. Y. (1989) *J. Bacteriol.* **171,** 4146–4153
21. Ye, Z. H., and Lee, C. Y. (1993) *J. Bacteriol.* **175,** 1095–1102
22. Chai, S., Kruft, V., and Alonso, J. C. (1994) *Virology* **202,** 930–939
23. McDonnell, G., Wood, H., Devine, K., and McConnell, D. (1994) *J. Bacteriol.* **176,** 5820–5830
24. Katz, M. E., Howarth, P. M., Yong, W. K., Riffkin, G. G., Depiazzi, L. J., and Rood, J. I. (1991) *J. Gen. Microbiol.* **137,** 2117–2124
25. Cheetham, B. F., and Katz, M. E. (1995) *Mol. Microbiol.* **18,** 201–208
26. Cheetham, B. F., Tattersall, D. B., Bloomfield, G. A., Rood, J. I., and Katz, M. E. (1995) *Gene* (*Amst.*) **162,** 53–58
27. Bruttin, A., and Brussow, H. (1996) *Virology* **219,** 96–104
28. Bruttin, A., Desiere, F., Lucchini, S., Foley, S., and Brussow, H. (1997) *Virology* **233,** 136–148
29. Ladero, V., Garcia, P., Bascaran, V., Herrero, M., Alvarez, M. A., and Suarez, J. E. (1998) *J. Bacteriol.* **180,** 3474–3476
30. Loessner, M. J., Inman, R. B., Lauer, P., and Calendar, R. (2000) *Mol. Microbiol.* **35,** 324–340
31. Narita, S., Kaneko, J., Chiba, J., Piemont, Y., Jarraud, S., Etienne, J., and Kamio, Y. (2001) *Gene* (*Amst.*) **268,** 195–206
32. Kass, E. H., Kendrick, M. I., Tsai, Y. C., and Parsonnet, J. (1987) *J. Infect. Dis.* **155,** 812–815
33. Todd, J. K., Todd, B. H., Franco-Buff, A., Smith, C. M., and Lawellin, D. W. (1987) *J. Infect. Dis.* **155,** 673–681
34. Wong, A. C., and Bergdoll, M. S. (1990) *Infect. Immun.* **58,** 1026–1029
35. Ross, R. A., and Onderdonk, A. B. (2000) *Infect. Immun.* **68,** 5205–5209
36. Yarwood, J. M., McCormick, J. K., Paustian, M. L., Kapur, V., and Schlievert, P. M. (2002) *J. Bacteriol.* **184,** 1095–1101
37. Daniels, D. L., Schroeder, J. L., Sanger, F., and Blattner, F. R. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., eds), p. 473, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
38. Kawabe, Y., and Ochi, A. (1990) *J. Exp. Med.* **172,** 1065–1070
39. Kawabe, Y., and Ochi, A. (1991) *Nature* **349,** 245–248
40. Attinger, A., Acha-Orbea, H., and MacDonald, H. R. (2000) *J. Immunol.* **165,** 1171–1174
41. Stolz, S. J., Davis, J. P., Vergeront, J. M., Crass, B. A., Chesney, P. J., Wand, P. J., and Bergdoll, M. S. (1985) *J. Infect. Dis.* **151,** 883–889
42. Schlievert, P. M. (1983) *J. Infect. Dis.* **147,** 391–398
43. Poindexter, N. J., and Schlievert, P. M. (1986) *J. Infect. Dis.* **153,** 772–779
44. Hofer, M. F., Newell, K., Duke, R. C., Schlievert, P. M., Freed, J. H., and Leung, D. Y. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5425–5430
45. Lee, P. K., Vercellotti, G. M., Deringer, J. R., and Schlievert, P. M. (1991) *J. Infect. Dis.* **164,** 711–719
46. Fast, D. J., Schlievert, P. M., and Nelson, R. D. (1988) *J. Immunol.* **140,** 949–953
47. Fast, D. J., Schlievert, P. M., and Nelson, R. D. (1989) *Infect. Immun.* **57,** 291–294
48. Schlievert, P. M., Assimacopoulos, A. P., and Cleary, P. P. (1996) *J. Lab. Clin. Med.* **127,** 13–22
49. Vergeront, J. M., Stolz, S. J., Crass, B. A., Nelson, D. B., Davis, J. P., and Bergdoll, M. S. (1983) *J. Infect. Dis.* **148,** 692–698
50. Childs, C., Edwards-Jones, V., Heathcote, D. M., Dawson, M., and Davenport, P. J. (1994) *Burns* **20,** 514–521
51. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1998) in *2-D Proteome Analysis Protocols* (Link, A. J., ed) pp. 531–552, Humana Press, Totowa, NJ